

# **Porovnání vybraných softwarových systémů pro tvorbu digitálních knihoven**

z oblasti Open Source pro potřeby NUŠL STK

Tento dokument obsahuje stručné charakteristiky a orientační porovnání softwarových systémů pro tvorbu digitálních knihoven. Vznikl pro Státní technickou knihovnu (STK) v rámci projektu Národní úložiště šedé literatury (NUŠL). Popisuje a porovnává systémy CDS Invenio, DSpace, Eprints, Fedora a Greenstone. Tyto systémy byly vybrány víceméně intuitivně na základě jejich rozšíření a známosti ve světě i v České republice. Všechny pocházejí z oblasti Open Source - z oblasti s otevřeným zdrojovým kódem. Tato otevřenost při dodržení určitých podmínek znamená technickou i legální dostupnost zdrojového kódu pro uživatele a tedy ekonomicky výhodnou možnost získat kvalitní software. Open Source ovšem neznamená software zdarma, i zde je třeba při výběru hledět na určitou stabilitu vývoje, dostupnost služeb podpory a rozvoje software atd. Tento dokument pouze shrnuje základní vlastnosti vybraných systémů a porovnává je na poměrně vysoké úrovni abstrakce, rozhodně nenahrazuje výběrové řízení softwarového systému pro NUŠL, vznikl pouze pro informativní podporu tohoto výběrového řízení.

## **CDS Invenio**

CDS Invenio, (dřívější název CDSWare), je vysoce modulární systém pro uchování a zpřístupnění velkého počtu knihovnických záznamů. Vznikl a dále se vyvíjí ve švýcarském CERNu. Vývoj začal v roce 1993 primárně pro interní potřeby této proslulé vědecké instituce jako institucionální repozitory, od roku 2000 zahrnuje podporu multimédií, OAI-PMH atd. a šíří se pod licencí GNU GPL i mimo CERN. Nyní představuje vyzrálé řešení, vyvíjené ve spolupráci s École Polytechnique Fédérale de Lausanne (EPFL), pro všeobecnou správu dokumentů, institucionální repozitory nebo rozsáhlý knihovní systém. Je lokalizován do 18 jazyků včetně češtiny, ruštiny ... Umožňuje vlastní definici matadatového schématu, interní jmenná konvence je podle MARC21. Podporuje standard OAI-PMH.

Systém CDS Invenio je velice propracovaný a flexibilní, lze jej nastavit mnoha různými způsoby pro nejrůznější použití, platí se za to ovšem jeho relativní složitostí. Ze systémů zahrnutých do tohoto hodnocení má ovšem nejvíce formalizovanou podporu při instalaci i provozu, poskytovanou za úplatu na základě smlouvy, (úrovni služeb a podpory se mu blíží systém Eprints). Kromě systému samotného lze počítat i s lokalizací některých služeb do češtiny, případně slovenštiny.

CDS Invenio se skládá z mnoha více či méně nezávislých modulů s přesně definovanou funkcionalitou. Následuje jejich stručný popis, z kterého jsou zřejmé další vlastnosti systému.

**BibCheck** je nástroj pro administrátora a správce obsahu pro automatizaci různých testů kvality metadat jako jsou integritní pravidla, formáty a rozsahy polí atd.

**BibClassify** automaticky extrahuje klíčová slova z dokumentů na základě frekvence výrazů obsažených ve slovníku (thesaurus).

**BibConvert** umožňuje konverzi metadat z jakýchkoliv strukturovaných formátů do jiných formátů. Pracuje podobně jako procesor XSLT. Vstupní a výstupní formáty jsou plně konfigurovatelné, popisy běžných formátů jsou obsaženy v dodávce.

**BibEdit** umožňuje editaci metadat prostřednictvím webového rozhraní.

**BibFormat** umožňuje formátovat bibliografická data mnoha různými způsoby. Odděluje administraci obsahu od formátování výstupů, může pracovat na pozadí v reálném čase, nebo může předformátovat často používané výstupy.

**BibHarvest** je harvester kompatibilní s [OAI-PMH](#), umožňuje získávat metadata z jiných úložišť podporujících OAI-PMH a správu úložiště. To je vytvořeno přímo v databázi a zahrnuje správcovský modul, který umožňuje administraci úložiště nezávisle na hlavní správě dat. Úložiště lze zcela nebo částečně otevřít pro harvesting v rozsahu protokolu OAI-PMH. Data jsou poskytována v nezpracované formě, jednotlivé tagy jsou pojmenovány podle konvence MARC21.

**BibIndex** indexuje metadata, odkazy (reference) a textové soubory. Vytváří dva druhy indexů – slovní a frázové. Lze definovat celou řadu logických indexů (autor, název atd.) Indexy se skládají ze dvou částí – dopředný a zpětný index. Toto indexování umožňuje velice rychlé prohledávání.

**BibMatch** filtruje vstupní soubory XML proti obsahu databáze a hledá záznamy podle různých kritérií, např. pro zamezení duplicitních záznamů.

**BibRank** umožňuje nastavit různá třídící/řadící kritéria pro pozdější použití při vyhledávání, např. podle frekvence výskytu určitého slova, podle počtu stažení dokumentu atd. BibRank je nezávislý na modulu BibIndex.

**BibSched** je časovač úloh, který jako centrální modul celého systému umožňuje ostatním modulům přístup k databázi kontrolovaným způsobem, zamezuje tak střetům při přístupu k databázi a zajišťuje konzistenci při aktualizaci databáze. Umožňuje administrátorovi monitorovat a řídit frontu úloh.

**BibUpload** umožňuje ukládat nová data do databáze. Vstup musí být dobře formátovaný soubor XML, zohledňující zvolené metadatové schéma. Ve většině případů jsou tato vstupní data výsledkem činnosti modulu BibConvert.

**ElmSubmit** je brána pro automatické nahrávání dokumentů z ověřených zdrojů prostřednictvím e-mailu. Obvykle se upřednostňuje poskytování souborů nebo harvesting prostřednictvím www.

**MiscUtil** je sada příležitostných nástrojů, které používají jiné moduly.

**WebAccess** je modul, který zajišťuje přístup uživatelů k systému. Využívá se technika RBAC (Role-Based Access Control – přístup podle rolí), kdy se uživatelé dělí do několika skupin podle jejich rolí a oprávnění v systému. Každá skupina má povolené jen určité operace v rámci systému.

**WebAlert** zajišťuje upozorňování uživatelů systému na události v systému, např. na výskyt nového dokumentu, splňujícího určitá kritéria. Tato kritéria odpovídají typickým uživatelským dotazům. Upozornění mohou být uživatelům zaslána e-mailem, nebo uložena v jejich košíku.

**WebBasket** umožňuje uživatelům ukládat dokumenty do osobního košíku, podobně jako se v internetovém obchodě ukládá zboží. Jeden uživatel může mít více košíků. Košík může být osobní nebo veřejný. Veřejný košík umožňuje sdílení dokumentů ve skupině uživatelů.

**WebComment** je skupinový nástroj pro hodnocení dokumentů a pro sdílení komentářů k dokumentům mezi čtenáři. V integraci se skupinovými nástroji WebBasket, WebGroup, WebMessage tools a WebComment je středem nástrojů pro skupinovou komunikaci v systému CDS Invenio.

**WebHelp** reprezentuje globální dokumentaci systému CDS Invenio pro uživatele a administrátory. Dokumentace specifická pro jednotlivé moduly je zahrnuta v jednotlivých modulech.

**WebMessage** umožňuje komunikaci mezi (eventuelně anonymními) uživateli systému prostřednictvím webových nástěnek, např. pro pozvánky do diskusních skupin atd.

**WebSearch** je modul, prostřednictvím kterého uživatelé zadávají dotazy do systému pro vyhledání určitých slov nebo frází v databázi. Umožňuje dva druhy dotazů: vyhledávání slov nebo frází. Dále umožňuje komplexní booleovské dotazy, vyhledávání pomocí regulárních výrazů, kombinované vyhledávání v metadatech a ve vlastních souborech. Uživatelé mohou záznamy prohlížet také podle indexů. Pokud je dotaz neúspěšný, systém nabízí alternativy a pomoc při vyhledávání. Metadata ve sbírkách jsou přímo přístupná k prohlížení podobně jako v populárním konceptu *Web Directories*. Ortogonální pohledy na dokumenty jsou organizovány do virtuálních sbírek, dokumenty lze klasifikovat podle více kritérií. Tato flexibilní organizace umožňuje vytvořit velice přehledná navigační schémata pro uživatele

**WebSession** je manažer relací a uživatelů, umožňuje rozlišovat uživatele a personalizovat uživatelské rozhraní jako jsou osobní košíky atd.

**WebStat** je konfigurovatelný systém, který umožňuje sběr údajů a statistik o systému jako takovém i o jeho užívání.

**WebStyle** je knihovna modulů, které definují design a vzhled systému CDS Invenio.

**WebSubmit** je komplexní systém, který autorizovaným uživatelům umožňuje ukládat do systému jednotlivé dokumenty. Systém poskytování obsahu zahrnuje kontrolní mechanismy, konverze různých formátů dokumentů a extrakci bibliografických údajů.

Více informací najdete na adrese <http://cdsware.cern.ch/invenio>

## **DSpace**

DSpace pochází jako většina systémů pro digitální knihovny z oblasti Open Source z univerzitního prostředí, vývoj byl zahájen knihovnou Massachusetts Institute of Technology (MIT) a společností Hewlett-Packard, první verze pochází z roku 2002. Nyní na projektu spolupracují další univerzity a jiné organizace, celkem cca 500 organizací v 60 zemích. Vývoj zastřešuje nezisková organizace DSpace Foundation, Massachusetts. Vývoj a hlášení chyb je technologicky zajištěno prostřednictvím SourceForge (sourceforge.net), pro vývoj jsou stanovena jasná pravidla, novými funkcemi může přispívat každý člen komunity.

DSpace je software určený pro tvorbu digitálních úložišť různých typů dokumentů pro nejrozličnější typy organizací. V současné době podporuje metadatové schéma Dublin Core, pro určení perzistentních identifikátorů používá CNRI Handle System, podporuje OAI-PMH 2.0 jako poskytovatel dat. Podporuje také standard OpenURL, umožňuje export dat do jednoduchého formátu XML nebo do formátu METS.

Data jsou v systému uložena v digitálních objektech s jednoznačným identifikátorem, které obsahují data i metadata. Dále existují v systému i struktury pro další logické členění záznamů do kolekcí a sbírek. Objekt může patřit do více logických struktur. DSpace podporuje tvorbu uživatelských účtů, seskupování uživatelů do skupin, přidělování přístupových práv a základní systém autentizace uživatelů.

Vyhledávat lze v systému DSpace prostřednictvím metadat i v samotných souborech, systém indexuje nejen prosté textové soubory, ale i některé další formáty, jako jsou PDF a Microsoft Word.

Poměrně kvalitní uživatelské webové rozhraní je tvořeno technologií kaskádových stylů a lze je tedy relativně snadno modifikovat. Poměrně čerstvou novinkou je systém Manakin, druhá verze uživatelského rozhraní systému DSpace, kde lze při modifikaci vzhledu pracovat se šablonami. Uživatelské rozhraní umožňuje všechny základní funkce, prohlížení a vyhledávání záznamů pro uživatele a ukládání záznamů pro správce obsahu. Na úrovni vkládání obsahu lze definovat pracovní postupy, které zahrnují několik úrovní schvalování obsahu před jeho vlastním uložením do databáze.

Systém DSpace má poměrně širokou a živou komunitu uživatelů a lze jej relativně snadno zprovoznit, má však omezenou funkcionalitu, modifikace je nutno většinou programovat, má také specifické nároky na provozní systém – UNIX a řada provozních komponent.

Více informací najdete na adresách

<http://www.dspace.org> a <http://www.ics.muni.cz/dspacecz>

## EPrints

Eprints je komplexní systém s mnoha instalacemi po celém světě pocházející opět z univerzitního prostředí, je vyvíjen v University of Southampton ve velké Británii, k dispozici je pod licencí GNU. Celý systém je založen na webových technologiích, primárním účelem je tvorba institucionálních úložišť mnoha typů dokumentů od běžné literatury přes vědecké a vysokoškolské práce po nejrozličnější jiné typy dokumentů včetně multimédií s primárním zaměřením na vědecké informace. Přestože celý systém je značně konfigurovatelný a lze při implementaci vytvořit obecné úložiště, jeho primární orientace na vědeckou oblast může takové úsilí komplikovat a omezovat. Při instalaci a provozu lze využívat placené služby, jako jsou školení, řízení implementačního projektu, technická podpora atd. – vše v angličtině.

V oblasti standardů podporuje systém EPrints OAI-PMH, metadata jsou reprezentována ve vlastním vnitřním formátu. Systém umožňuje import dat z formátu XML a z některých externích zdrojů (PubMed ...), export je možný v mnoha formátech, XML, RSS, DublinCore, METS.

Vnitřní architektura nepoužívá pojem digitální objekt, ale Item (položka). Položka v sobě zapouzdřuje metadata i vlastní soubory. V rámci jedné instalace lze vytvořit více úložišť i pro odlišné organizace, případně více oddělených sbírek pro jednu organizaci s odlišným vzhledem i strukturou. Technicky je to zajištěno provozem více instancí webového serveru Apache.

Systém Eprints podporuje správu uživatelských účtů, přidělování práv je méně propracované (v první řadě se počítá s publikujícími vědci). Systém rozlišuje tři typy uživatelských účtů – administrátor, editor a běžný uživatel. I běžný uživatel může do systému vkládat data, a to velice podrobně, zejména na úrovni metadat. Konečné uložení dat musí schválit uživatel s profilem editor.

Systém EPrints má podobně jako systém DSpace komplexní uživatelské rozhraní, založené na webových technologiích. Rozhraní lze poměrně dobře konfigurovat, část rozhraní tvoří statické stránky v HTML, které systém generuje pro vložené položky – nepočítá se s denním vkládáním velkého počtu záznamů.

Prostřednictvím rozhraní lze vyhledávat v metadatech i v datech, kromě textových souborů systém indexuje i některé jiné běžné formáty (PDF), obsah lze procházet i v logických stromových strukturách, výchozí struktura je stejná jako v Library of Congress, USA, lze ji však modifikovat. Rozhraní dále umožňuje registrovat nové uživatele, informovat se o novinkách v systému atd. Administrátorské rozhraní umožňuje celý systém konfigurovat a obsluhovat.

Celý systém je velice komplexní, vkládání položek je propracované, nicméně relativně složité a časově náročné – počítá se především s publikujícími vědci, kteří vkládají malé množství jednotek s velkým objemem informací ke každé. Tento rys může komplikovat tvorbu obecných úložišť.

Více informací najdete na adrese <http://www.eprints.org>

## Fedora

Systém Fedora (Flexible Extensible Digital Object Repository Architecture) opět pochází z univerzitního prostředí – Cornell University a University of Virginia. Vše začalo jako výzkumný projekt v roce 1997, jehož výsledek byl zpřístupněn na webu Cornell University v roce 1998. V roce 2001 se spojily obě university a získaly prostředky na další vývoj od Mellonovy nadace se zadáním vyvinout univerzální digitální knihovnu na základě webových služeb a XML. V roce 2007 obě university založily organizaci Fedora Commons, která se nyní stará o vývoj jejich společného systému.

Systém Fedora je digitálním úložištěm, které umožňuje ukládat a uchovávat v podstatě libovolné digitální objekty, nicméně nabízí právě jen služby úložiště ve formě jádra, které se stará o ukládání a správu digitálních objektů plus programátorské knihovny funkcí a volání služeb, které umožňují vývojářům obsluhovat úložiště. Systém Fedora nemá vlastní uživatelské rozhraní, to je nutné vytvořit pro každé nasazení systému zvlášť. Přestože je systém Fedora velice kvalitním úložištěm, respektujícím základní teoretické poznatky v oblasti budování digitálních knihoven, tak skutečnost, že systém není kompletní a je nutno při jeho nasazení doprogramovávat uživatelské rozhraní, brání jeho použití v projektech, kde je nutné rychlé nasazení, případně kde nejsou vývojové kapacity.

Základním prvkem architektury systému je podobně jako u jiných systémů digitální objekt, nicméně zde se nejedná pouze o objekt ve smyslu uloženého záznamu, ale systém rozlišuje tři druhy objektů:

- datový objekt, (Data Object), který obsahuje metadata a data, případně jednoznačný perzistentní identifikátor, ukazující na data v jiném úložišti, (systém umožňuje verzování - ukládání různých verzí téhož dokumentu),
- dále objekt popisu chování, (Behaviour Definition Object), popis služeb, které se váží k určitému souboru (například zobrazení dat určitého typu), tyto služby jsou popsány jazykem WDSL a zajišťují tak relativní platformovou nezávislost pro další typ objektu
- objekt mechanismu (implementace) chování, (Behaviour Mechanism Object), (např. software, který umožní data zobrazit), tyto služby mohou být napsány teoreticky v jakémkoliv programovacím jazyce.

Systém Fedora podporuje celou řadu standardů, OAI-PMH, export do formátů METS a vlastního interního formátu FOXML, popisná metadata jsou ukládána ve formátu Dublin Core.

Přestože je jádro systému Fedora velice pokrokové, nejedná se v současné době o plnohodnotný kompletní knihovní systém. Je to jen základ, který je potřeba dotvořit s poměrně velkými náklady a s velkým úsilím. I při provozu je nutno kvůli platformové nezávislosti systému třeba počítat s větší režii, je náročnější na hardwarové zdroje než ostatní systémy.

Více informací najdete na adrese <http://www.fedora.info>

## Greenstone

Systém Greenstone (Greenstone Digital Library Software) je exotičtější svým původem než jeho kolegové, nicméně opět se jedná o univerzitní produkt – je vyvíjen od roku 1995 na University of Waikato na Novém Zélandu, vývoj podporuje i organizace UNESCO. Je dostupný pod licencí GNU GPL. Přes svůj exotický původ je systém kromě jiných jazyků částečně lokalizován i do češtiny.

V oblasti standardů systém podporuje metadatové schéma Dublin Core, protokol OAI-PMH a jako jediný ze zkoumaných systémů podporuje protokol Z39.50 pro obousměrné prohledávání katalogů jiných knihoven, nicméně tato podpora není standardně nastavena a systém je nutné pro její zprovoznění překompilovat.

Základní jednotkou pro uložení v systému je dokument, (systém podporuje mnoho formátů digitálních objektů), který se ukládá do předem připravených sbírek, v nichž je definováno, jaké dokumenty budou obsahovat, a tyto sbírky se pak sdružují do knihoven. Pro zpracování jednotlivých typů dokumentů musí systém obsahovat programovou podporu, tzv. pluginy. Tato architektura klade větší nároky na nastavení systému, než je tomu u některých jiných kompletnějších systémů. Při uložení dokumentu do předem definované sbírky v interním formátu XML je vygenerován i jednoznačný identifikátor ve vlastním formátu.

Uživatelské rozhraní je oddělené od vlastního úložiště (je to separátní aplikace), úložiště i uživatelské rozhraní spolu komunikují vlastním interním protokolem, který lze nahradit protokolem CORBA a připojit tak k systému vlastní aplikaci. Uživatelské rozhraní lze nastavovat pro každou sbírku jinak, je založené na webových technologiích.

Prostřednictvím něj lze procházet sbírky a vyhledávat v nich. Procházení sbírek je založeno na řazení příbuzných dokumentů do společných složek, které lze hierarchicky propojovat a vytvářet tak logickou prohledávací strukturu.

Systém obsahuje správu uživatelů a přidělování uživatelských přístupových práv, ale nemá tuto podporu příliš dobře propracovanou. Omezuje se na přidělování práv administrátorům a správcům obsahu, nepočítá se s aktivitou jiných uživatelů.

Systém je zcela jistě životaschopný, jsou k němu dostupné i služby, poskytované organizací DL Consulting na Novém Zélandě, nicméně jeho nasazení je relativně složité a použití je omezeno na organizace, kde se nepředpokládá vkládání dat uživateli mimo správce sbírek.

Více informací najdete na adrese <http://www.greenstone.org>

## Přehledová tabulka

	CDS Invenio	DSpace	EPrints	Fedora	Greenstone
Rok vzniku	1993	2002	2000	1997	1997
Zajištěn vývoj	Ano	Ano	Ano	Ano	Ano
Vývojová organizace	CERN, Švýcarsko	DSpace Foundation, Massachusetts, USA (MIT/HP)	University of Southampton, Velká Británie	Cornell University, University of Virginia, USA	University of Waikato, Nový Zéland
Programovací jazyk	Python (PHP, Common LISP)	Java	PERL	Java	PERL, GDBM, MG
Operační systém, provoz	Unix, Python/Apache	Unix, Apache... (Windows/XP)	Unix, Apache, PHP ...	Unix, Apache	Unix, Linux, Windows ...
Databáze	MySQL	PostgreSQL Oracle	MySQL	MySQL, Oracle ...	Není třeba
OAI-PMH	Ano	Ano, poskytovatel	Ano	Ano	Ano
Z39.50	Ne	Ne	Ne	Ne	Ano
Metadatový formát	MARC21, flexibilní	Dublin Core	Dublin Core	Dublin Core, METS ...	Dublin Core
Identifikátory	Vlastní	CNRI Handle	Vlastní	Vlastní	Vlastní



## **Základní vyhodnocení systémů se zaměřilo na tyto oblasti:**

- **Řízení obsahu** – nástroje a postupy podporující ukládání obsahu do digitální knihovny a řízení procesu tohoto ukládání, verzování. V této kategorii neměl žádný ze systémů vážnější nedostatky, pouze Fedora postrádá formalizovanou podporu pro ukládání obsahu, na druhou stranu umožňuje ukládání verzí. Nejkomplexnější nástroje v této oblasti nabízí CDS Invenio.
- **Získávání obsahu** – import a export obsahu, podpora různých formátů dokumentů – v této kategorii opět nemá žádný systém výrazné slabiny, Eprints a CDS Invenio vychází v této kategorii jako systémy s nejkomplexnější podporou.
- **Metadata** – podpora různých metadatových formátů je důležitá pro indexování, ukládání, zpřístupňování a ochranu obsahu. Většina systémů v této kategorii podporuje pouze některé metadatové standardy, většinou Dublin Core a MARC21, Fedora nativně podporuje i některé méně rozšířené formáty (METS), CDS Invenio umožňuje mapovat jakýkoliv metadatový formát na jeho vnitřní strukturu, reprezentovanou jmennou konvencí MARC21.
- **Podpora vyhledávání** – týká se celé řady vyhledávacích a prohlížečích funkcí, vyhledávání v metadatech, fulltextové vyhledávání, hierarchické prohlížení apod. Toto je velice důležitá kategorie a všechny systémy nabízejí základní vyhledávací funkce. Fedora zde ztrácí body svou absencí hotového uživatelského rozhraní, CDS Invenio nabízí velice propracované vyhledávání i hierarchické procházení. Žádný ze systémů překvapivě nenabízí proximitní vyhledávání.
- **Řízení přístupu a ochrana soukromí** – řízení přístupu a ochrana soukromí zahrnuje administraci hesel, uživatelských účtů s přístupovými právy včetně možnosti získat zapomenuté heslo atd. Zde nabízí nejpropracovanější podporu CDS Invenio zejména ve správě uživatelských hesel s možností přidělování hesel systémem, vlastní volby hesla i s možností získat zapomenuté heslo. Nejméně propracovaný systém v této oblasti má systém Greenstone.
- **Podpora výstupů a dotazování** – toto kritérium souvisí s možností vyhodnocovat využití digitální knihovny a odhalovat vzorce chování uživatelů pro vylepšování poskytovaných služeb, zaznamenávání aktivit uživatelů pro účely účtování. Zde nabízí nejlepší podporu systém Greenstone, Fedora nabízí také statistiky o užívání, nicméně nenabízí možnost úpravy výstupních zpráv. Eprints a CDS Invenio nemají mnoho hotových výstupů. CDS Invenio zejména v modulu WebStat nabízí možnost tvorby výstupů o užívání systému.
- **Udržitelnost, ochrana dat** – ochrana metadat, konzistence a integrity datové základny, zálohování, podpora případné migrace dat. Všechny systémy nabízejí možnost migrace prostřednictvím exportu do různých formátů, zejména XML, Fedora má velice propracovanou strategii v této oblasti, DSPace jako jediný systém používá

nativně CNRI Handles, systém Greenstone používá perzistentní identifikátory vlastního formátu, systémy Fedora a CDS Invenio jsou připraveny pro spolupráci s jakýmkoliv systémem persistentních identifikátorů.

- **Interoperabilita** – možnost obousměrné integrace s jinými distribuovanými systémy na úrovni metadat, prohledávání i získávání a poskytování dokumentů, podpora OAI-PMH, Z39.50 - Z39.50 podporuje pouze systém Greenstone, OAI-PMH podporují v té či oné podobě všechny systémy, DSpace pouze jako poskytovatel dat.
- **Uživatelské rozhraní** – tato kategorie se týká podpory více jazyků a schopnosti přizpůsobit uživatelské rozhraní různým potřebám různých uživatelů či rozdílných implementací. V této kategorii je nejslabším systémem Fedora, protože má pouze jednoduché testovací rozhraní a vlastní rozhraní pro uživatele je třeba dotvořit. Skvělou koncepcí uživatelského rozhraní má systém DSpace – kaskádové styly a šablony Manakin, CDS Invenio nabízí velice dobré nástroje jak pro jednotlivé uživatele, tak pro skupinovou komunikaci. Českou lokalizaci nabízí CDS Invenio a částečně Greenstone, DSpace má velkou aktivní uživatelskou komunitu, která je v této oblasti také činná.
- **Podpora standardů** – standardy jsou důležité pro sdílení a dlouhodobé uchování digitálního obsahu. Jedná se o celé spektrum oblastí standardů od metadat přes interoperabilitu až po formáty uložených dokumentů. Viz popisy jednotlivých systémů – základní důležité standardy podporují všechny systémy, Z39.50 podporuje pouze systém Greenstone. Nativní podporu CHRI Handles má pouze systém DSpace.
- **Nástroje pro automatizaci** – tato kategorie se týká nástrojů pro automatizované získávání obsahu, harvesting, generování metadat, údržbových činností atd. V této oblasti vynikají systémy Greenstone a CDS Invenio, Fedora a Eprints nemají tak dobrou podporu automatizace, DSpace je někde mezi těmito dvěma skupinami.
- **Podpora, služby** – důležitý aspekt u všech softwarových systémů. Mnoho dobrých systémů digitálních knihoven se mj. nachází v oblasti Open Source, kde je třeba na toto hledisko dbát zejména. Důležitá je dokumentace, helpdesk, sběr požadavků na vylepšení, případně diskusní fóra atd. Formalizovanou placenou podporu na základě smlouvy mají jen systémy CDS Invenio a Eprints, systém Eprints má zřejmě největší instalovanou bázi ve světě, DSpace má velice živou uživatelskou komunitu a podporu prostřednictvím SourceForge.

Uvedené hodnocení je relativně subjektivní zejména vzhledem k omezené možnosti poznání některých systémů v České republice. V následující tabulce uvádíme přibližná číselná skóre jednotlivých systémů v uvedených kategoriích na stupnici od jedné do deseti.

	CDS Invenio	DSpace	Eprints	Fedora	Greenstone
Řízení obsahu	10	8	8	5	9
Získávání obsahu	10	9	10	8	9
Metadata	6	5	5	7	5
Podpora vyhledávání	8	6	5	4	7
Řízení přístupu	8	6	8	4	5
Podpora výstupů	6	6	4	6	10
Udržitelnost, ochrana dat	8	8	7	10	6
Interoperabilita	6	6	6	7	8
Uživatelské rozhraní	8	8	7	5	7
Podpora standardů	9	9	9	9	10
Nástroje pro automatizaci	9	7	5	5	8
Podpora, služby	9	7	8	6	6
Součet	97	85	82	76	90

Pokud hodnotíme uvedené systémy jen podle jejich vlastností bez ohledu na vhodnost nasazení v našich podmínkách, nejlépe vycházejí systémy Greenstone a CDS Invenio, těsně následovány systémem DSpace. Fedora není hotový systém a jeho dotvoření je pracné a nákladné, systém Eprints je poměrně jednoúčelový pro menší denní přírůstky důkladně popisovaných vědeckých dokumentů.

Pro nasazení v České republice je dostupností služeb a jazykové lokalizace vhodný systém CDS Invenio, který je zároveň velice funkčně bohatý a flexibilní, za což se ovšem platí jeho relativní implementační a provozní složitostí.

## Rešerše:

*Apache : HTTP server project* [online]. s.l. : The Apache Software Foundation, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://httpd.apache.org/>>.

Cattaneo, Gianpiero. *Authentication/ authorization issues and fulltext dokument migration for the CERN Document Server* [online]. s. l. : CERN, 2007. [cit. 2008-11-09]. Dostupný z WWW: <<http://cdsweb.cern.ch/record/1063371/files/thesis-twoside.pdf>>.

*CDS Software Consortium : CDS Invenio* [online]. s.l. : CERN, 2005-2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://cdsware.cern.ch/invenio>>.

*Common Object Request Broker Architecture (CORBA)* [online]. s.l. : Object management group, 1997-2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.corba.org/>>.

*Czech DSpace* [online]. s.l. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.ics.muni.cz/dspacecz/>>.

Davis, Jim et al. *Dienst Protocol Specification* [online]. Ithaka : Cornell University, 2000. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.cs.cornell.edu/cdlrg/dienst/protocols/DienstProtocol.htm>>.

*Digital Asset Assessment Tool (DAAT) Project : Risk factors : An assessment of the risk factors which may affect survival of digital assets* [online]. London : University of London Computer Centre, 2006. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.ulcc.ac.uk/daat.html>>.

Dijk, Elly. Sharing grey literature by using OA-x. *The Grey Journal*, 2005, vol 1. no 2, p. 83-90. ISSN 1574-1796.

*DSPACE* [online]. s.l. : DSpace, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.dspace.org>>.

*EPrints* [online]. s.l. : EPrints, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://software.eprints.org/>>.

*Fedora : Flexible Extensible Digital Object and Repository Architecture* [online]. Ithaka : Fedora Commons, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.fedora.info/>>.

*Greenstone Digital Library Software* [online]. s.l. : Greenstone, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.greenstone.org/>>.

*GNU Operating System* [online]. s.l. : Free Software Foundation, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.gnu.org/>>.

HALBICH, Ľestmír. Užití Open alternativ v knihovních a informačních službách. In *INFORUM 2007: 13. konference o profesionálních informačních zdrojích Praha, 22. - 24.5. 2007*. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.inforum.cz/pdf/2007/halbich-cestmir.pdf>>.

HALBICH, Č., The software for Open Document Format in agrarian sector, (in Czech language), In *Proceedings of Agrarian perspectives XV. - Foreign trade and globalisation processes*, Prague 2006, ISBN 80-213-1531-8.

*Handle System* [online]. s. l. : CNRI, Corporation for National Research Initiatives, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.handle.net/>>.

*i-Tor OAI and OA-X: Tools and Technology for Open Repositories* [online]. [Amsterdam] : Netherlands Institute for Scientific Information Services, 2004. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.i-Tor.org>>.

Jalůvka, Dušan. *Úložiště digitálních dat pro potřeby ÚK VŠB TU Ostrava*. Ostrava, 2006. 61 s., příl. CD. Diplomová práce. Vysoká škola báňská - Technická univerzita Ostrava, Fakulta elektrotechniky a informatiky. Katedra informatiky. Konzultant diplomové práce Petr Šaloun.

Krejčíř, Vlastimil. Systémy pro tvorbu digitálních knihoven. In *INFORUM 2006: 12. konference o profesionálních informacích zdrojích Praha, 23. – 25. 5. 2006*. [cit. 2008-11-09]. Dostupný z WWW: <[http://www.inforum.cz/pdf/2006/Krejcir\\_Vlastimil.pdf](http://www.inforum.cz/pdf/2006/Krejcir_Vlastimil.pdf)>.

Krejčíř, Vlastimil. *Univerzální digitální repozitář*. [online]. Brno, 2005. Diplomová práce. Masarykova univerzita, Fakulta informatiky. [cit. 2008-11-09]. Dostupný z WWW: <<http://eprints.rclis.org/archive/00005076/>>.

Lagoze, Carl; Sompel van de, Herbert. Open Archives Initiative Object Reuse and Exchange. Ithaka. : s. n, c2008. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.openarchives.org/ore/documents/oreBetaPressRelease.pdf>>.

Louisiana Coastal Wetlands and Louisiana Coastal Grey Literature: Vanishing Treasures. *Publishig Research Quarterly*. 2007, vol 23, s. 19-29.

McGILL, Michael; RACINE, Drew. OCLC : Online Computer Library Center. In ARMS, Caroline (ed.). *Campus strategies for libraries and electronic information*. Bedford (Mass.) : Digital Press, 1990, s. 36-56.

Nixon, William. DAEDALUS: Initial experiences with EPrints and DSpace at the University of Glasgow [online]. *Ariadne*, 2003, issue 37. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.ariadne.ac.uk/issue37/nixon/intro.html>>.

*Open Archives Initiative Protocol for Metadata Harvesting* [online]. s.l. : OAI-PMH, c2004. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.openarchives.org/OAI/openarchivesprotocol.html>>.

*Open Source Software Trials in Government : Final report*. Norfolk : Office of Government Commerce, 2004. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.ogc.gov.uk/documents/CP0041OpenSourceSoftwareTrialReport.pdf>>.

Pepe, A. et al. *CERN Document Server Software: the integrated digital library* [online]. Geneva : CERN, [2006]. [cit. 2008-11-09]. Dostupný z WWW: <<http://cdsware.cern.ch/invenio/doc/elpub2005.pdf>>.

Příbramská, Iva; Vojna, Martin. *Návod na zprovoznění Z39.50* [online]. Praha : SUAleph, c1996-2001. [cit. 2008-11-09]. Dostupný z WWW: <<http://www.sualeph.cz/z39.htm>>.

Revere, Debra ; Bugni, Paul F.; Fuller, Sherrilynne. A Public Health Knowledge Management Repository that Includes Grey Literature. In *GL 8 international conference on grey literature*. Amsterdam : GreyNet, 2006.<[http://www.greynet.org/images/GL8\\_Revere\\_et\\_al.ppt](http://www.greynet.org/images/GL8_Revere_et_al.ppt)>.